Angle-Closure Detection in Anterior Segment OCT based on Multi-Level Deep Network

Huazhu Fu*, Yanwu Xu*, Stephen Lin, Damon Wing Kee Wong, Mani Baskaran, Meenakshi Mahesh, Tin Aung, and Jiang Liu

Abstract—**Irreversible visual impairment is often caused by primary angle-closure glaucoma, which could be detected via Anterior Segment Optical Coherence Tomography (AS-OCT). In this paper, an automated system based on deep learning is presented for angle-closure detection in AS-OCT images. Our system learns a discriminative representation from training data that captures subtle visual cues not modeled by handcrafted features. A Multi-Level Deep Network (MLDN) is proposed to formulate this learning, which utilizes three particular AS-OCT regions based on clinical priors: the global anterior segment structure, local iris region, and anterior chamber angle (ACA) patch. In our method, a sliding window based detector is designed to localize the ACA region, which addresses ACA detection as a regression task. Then, three parallel sub-networks are applied to extract AS-OCT representations for the global image and at clinically-relevant local regions. Finally, the extracted deep features of these sub-networks are concatenated into one fully connected layer to predict the angle-closure detection result. In the experiments, our system is shown to surpass previous detection methods and other deep learning systems on two clinical AS-OCT datasets.**

Index Terms—**AS-OCT, deep learning, angle-closure detection, anterior chamber angle.**

I. INTRODUCTION

The foremost cause of irreversible blindness is glaucoma, with primary angle-closure glaucoma (PACG) being a major cause of blindness in Asia [1], [2]. Since vision loss from PACG cannot be undone, it is essential to screen people at early asymptomatic stages of the disease to prevent vision loss. In general, gonioscopy is a convenient tool for simple opportunistic case detection in clinics, but imaging is envisaged to reduce this burden of screening for angle closure. Recently, anterior segment optical coherence tomography (AS-OCT)

*Corresponding authors: H. Fu and Y. Xu.

H. Fu is with Inception Institute of Artificial Intelligence, Abu Dhabi, United Arab Emirates, and also with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632 (e-mail: huazhufu@gmail.com).

Y. Xu is with the Department of Artificial Intelligence Innovation Business, Baidu Inc., Beijing 100193, China (e-mail: ywxu@ieee.org).

S. Lin is with Microsoft Research, Beijing 100080, China.

D. Wong is with Singapore Eye Research Institute, Singapore 168751, and also with Nanyang Technological University, Singapore 138632.

M. Baskaran and M. Mahesh are with the Singapore Eye Research Institute, Singapore 168751.

T. Aung is with Singapore Eye Research Institute, Singapore 168751, and also with the Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119077.

J. Liu is with the Cixi Institute of Biomedical Engineering, Chinese Academy of Sciences, Zhejiang 315201, China, and also with the Department of Computer Science and Engineering, Southern University of Science and Technology, Guangdong 518055, China.
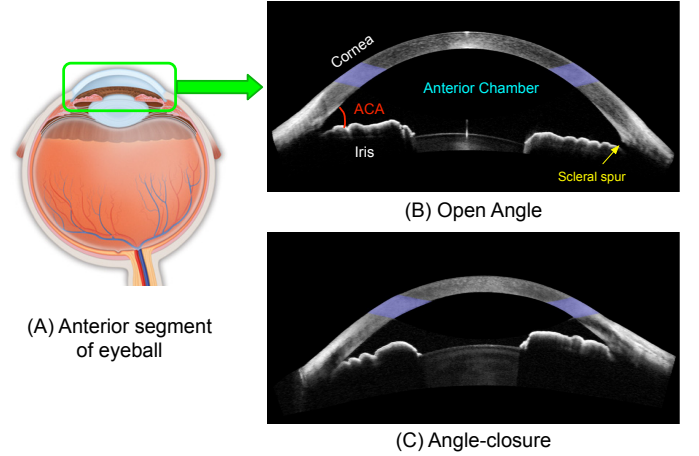

Fig. 1. (A) AS-OCT image captures the anterior segment region of the eye, including cornea, iris, ciliary body and lens. (B) Open angle. (C) Angle-closure.

technology has enabled high-resolution imaging of anterior segment structures [3], [4]. In contrast to fundus imaging, AS-OCT imaging provides clear cross-sections of the entire anterior chamber for observation. In a specific type of angle-closure, the anterior chamber angle (ACA) between the iris and cornea is narrow and leads to blockage of drainage channels as in Fig. 1 (C). This blockage can result in rising eye pressure and optic nerve damage [5]. Previous studies have suggested that several anterior chamber measurements such as the anterior chamber angle, width and lens vault are associated with angle closure [6], [7]. When clinicians use AS-OCT to image the angles for diagnosing angle closure in their patients, the AS-OCT images require a degree of interpretation. Moreover, AS-OCT imaging has limitations when used in clinical management, because of the lack of automated methods to interpret AS-OCT images for the presence of angle closure.

Several automated assessment methods for wide-scale angle-closure detection using AS-OCT have been studied [8]–[12]. A semi-automated system is provided in [13] to produce various anterior segment clinical measurements, but it needs the user to identify the positions of the Scleral Spurs, making it unsuitable for automated analysis of larger-scale datasets. Among fully automated systems, Tian *et al.* provide a clinical parameter calculation method for High-Definition OCT (HD-OCT) based on detecting the Schwalbe's line [8]. Xu *et al.* employ visual features directly obtained from the AS-OCT images to classify the glaucoma subtype with a support

vector machine (SVM) [9], [10]. In [11], [14], a label transfer method is proposed to combine segmentation, measurement and detection of AS-OCT structures. However, these methods mainly depend on segmentation to extract handcrafted representations, and then conduct angle-closure detection. Recently, deep learning techniques have been shown to surpass the performance of handcrafted features in many areas [15]–[17]. For example, Convolutional Neural Networks (CNNs) have led to improved performance in image classification [18], object detection [19] and image segmentation [20]. For retinal fundus images, there have also been several recent works achieving significant performance gains for diabetic retinopathy [21]–[23], age-related macular degeneration [24], [25] and glaucoma [26], [27]. These works inspire our examination of deep learning for angle-closure detection in AS-OCT images, and how clinical priors can be utilized to guide development of a suitable deep learning architecture.

In this paper, a deep learning system is proposed for angle-closure detection that makes use of two clinical observations: 1) the main clinical parameters for angle-closure are measured from different regions of the cornea, iris and anterior chamber, and are evident at different scales [6], [28]; 2) the intensities of AS-OCT images could vary among different AS-OCT imaging devices, which could degrade feature extraction and classification. Based on these clinical priors, we introduce a Multi-Level Deep Network (MLDN). In our method, multiple parallel sub-networks jointly learn multi-level representations from the multiple regions/levels known to be informative for angle-closure detection in an AS-OCT image. Furthermore, we also propose an intensity-based data augmentation method to elevate robustness to different AS-OCT imaging modalities. Our contributions can be summarized as follows:

1) We propose an automated angle-closure detection system by introducing a deep learning algorithm for the purpose of obtaining predictive representations in AS-OCT images.
2) A multi-level deep architecture is developed to account for clinical priors at different image areas and levels, in which multi-level features are extracted on different image levels through separate and parallel deep sub-networks.
3) An efficient and simple ACA detection method is utilized based on sliding-window regression, which formulates the ACA localization task as a regression problem.
4) Intensity-based data augmentation is presented to effectively handle intensity variations among the different AS-OCT imaging modalities.
5) Two clinical AS-OCT datasets are collected for evaluation, including the Visante AS-OCT and Cirrus HD-OCT modalities. Experiments show that the proposed MLDN outperforms the previous assessment methods and other deep learning architectures.

The rest of the paper is organized as follows: Sec. II reviews some related works. Sec. III introduces the proposed MLDN method and its main components. Sec. IV presents an evaluation of our method on two AS-OCT datasets, and Sec. V concludes the paper.

II. RELATED WORK

A. Angle-closure detection in AS-OCT image

The angle-closure detection systems in AS-OCT images can be categorized into two classes. The first is clinical quantitative parameter based, which calculates clinically-relevant measurements based on segmentation results for classification into glaucoma types [8], [11], [29], [30]. The clinical quantitative parameters are predefined based on anatomical structures [6], [28], [31], [32] (e.g., anterior chamber width, lens-vault, and trabecular-iris angle), which have explicit clinical meaning used for reference in making a diagnosis. Although such parameters are helpful to clinicians, their measurement depends heavily on the performance of AS-OCT segmentation. A different kind of screening method directly extracts visual features from the AS-OCT images [9], [10], [33]. Visual representations could mine a wider set of image information, and can obtain better screening performance than from clinical parameters. For example, the works in [9], [10] employ the Histogram of Oriented Gradients (HOG) [34] and histogram equalized pixel (HEP) [35] with linear SVM to represent the ACA region. An Anterior Chamber Shape feature is proposed in [33] to describes iris shape information. However, hand-crafted visual features such as these may not provide sufficiently discriminative representations. In our work, we present a method based on visual features, but learn a rich hierarchical representation for angle-closure detection in AS-OCT through deep learning.

B. Deep learning

Based on advances in discriminative representations and large scale data analysis, deep learning techniques have been demonstrated to achieve significant performance gains in many areas, e.g., image classification [18], [36], object detection [37]–[39], image segmentation [20], [40], [41], and person re-identification [42], [43]. For learning a rich hierarchical representation, the concept of multi-scale and multi-region networks have been proposed. For example, the multi-column network [44], [45] has different receptive field sizes for each sub-network, and combines the output maps, as shown in Fig. 2 (A). The skip-layer network [46], [47] combines features from different pooling layers to build a representation, as shown in Fig. 2 (B). Another deep architecture is the multi-region network [37], [48], which distinguishes object regions/proposals by using individual networks, as shown in Fig. 2 (C). The target regions are selected by using general object detectors [49], [50]. Our multi-level network extracts image representations at different levels, but generates a more concise feature focusing both on global and local clinically-relevant regions. The clinical guidance is introduced to extract appropriate regions for angle-closure detection, as shown in Fig. 2 (D). Avoiding redundant information in this manner facilitates deep network training.

III. PROPOSED METHOD

In this paper, a Multi-Level Deep Network (MLDN) for automated angle-closure detection in AS-OCT images is proposed. As shown in Fig. 3, in our system, the ACA region
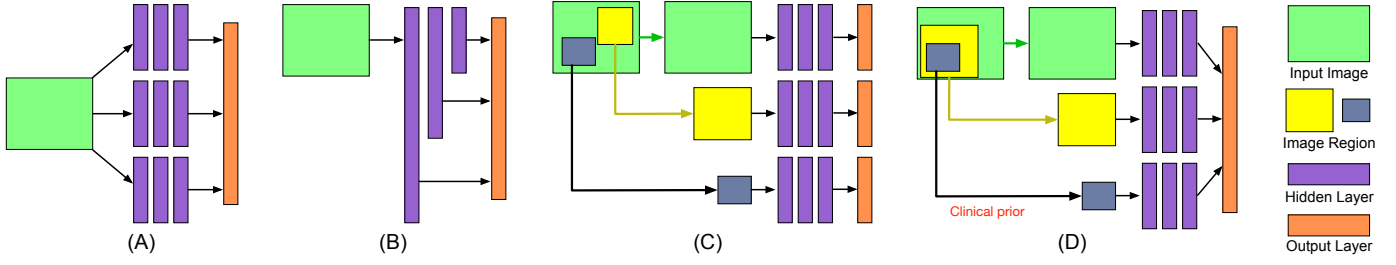
Fig. 2. Different deep learning architectural configurations: (A) multi-column architecture; (B) multi-scale architecture; (C) multi-region architecture; (D) our proposed multi-level architecture, where regions are extracted based on clinical guidance. The green block denotes the global image, while the yellow and blue blocks are proposed regions of attention. The purple and orange blocks denote the hidden layers for deep learning (e.g., convolutional layer) and the final output layer (e.g., fully connected layer), respectively.
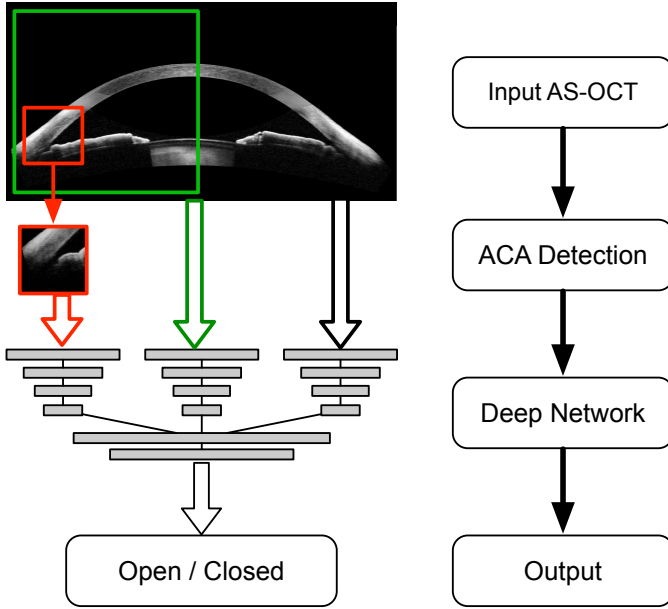

Fig. 3. System flowchart.

is detected by using sliding windows, and then the proposed MLDN combines global and local level representations for angle-closure detection in AS-OCT images. Our MLDN consists of three parallel sub-networks to extract hierarchical representations from different clinically-relevant regions in an AS-OCT image, including the global anterior segment structure, local iris region, and ACA patch. Finally, the output maps of the three sub-networks are concatenated as input into one fully connected layer to predict the angle-closure detection result.

A. ACA region detection

The Anterior Chamber Angle (ACA) lies at the junction of the cornea and the iris, as shown in Fig. 1 (B), and is a major bio-marker in AS-OCT images, providing key risk indicators for angle closure. Several methods have been proposed to detect the ACA region. For example, the work in [9] proposed a geometric method based on edge detection and shape properties to detect the ACA region. But this method only works well for certain AS-OCT imagery, and the

quality of image will greatly affect performance. Tian *et al.* provided an algorithm to localize the ACA for High-Definition OCT (HD-OCT) images by using the Schwalbe's line [8]. However, it is limited to HD-OCT images and not valid for low resolution modalities like AS-OCT. An automated AS-OCT structure segmentation based on label transfer was presented in [11], which transfers manually marked labels to the target image for guiding segmentation. In our work, we only require localization of the ACA region without pixel-level segmentation. We instead employ a simple but effective method based on a sliding window regression framework. In our method, we first segment the corneal boundary, and then a series of ROIs are extracted along the corneal boundary as regression candidates. The aim is to detect the Scleral Spur (SS) as a landmark, and crop a small region of interest (ROI) centered on the detected SS as the ACA patch.

Given the AS-OCT image, we first segment the corneal boundary, which is the top curve of the anterior segment region, as shown in Fig. 4 (A). After reducing speckle noise with a 2-D Gaussian filter, we calculate the vertical gradient map of the AS-OCT image, as shown in Fig. 4 (B). The corneal upper and bottom boundaries appear at the top negative and positive impulse responses in each column of the gradient map, respectively, as shown in Fig. 4 (C). We estimate the initial points by finding the top two edge points for each column in the AS-OCT, as shown in Fig. 4 (D), and then two fourth-order polynomials [11] are fitted to estimate the corneal boundary as shown in Fig. 4 (E).

With the detected corneal boundary, we sample a series of ROIs (with a size of 120×120 in our paper) as candidates along the corneal bottom boundary to localize the SS point, which is formulated as a regression task, as shown in Fig. 4 (F). For regression, the distance between the ROI center to the SS point is used:

$$d = \min \left\{ 1, \frac{2|v_r - v_s|}{W_r} \right\}, \tag{1}$$

where the v_r and v_s denote the horizontal coordinates of the ROI center and SS point, respectively. W_r is the width of the ROI. The $\min\{1, *\}$ ensures that the distance lies in the range of $[0, 1]$ when the SS point is outside of the ROI. In our method, we employ L2-regularized support vector regression
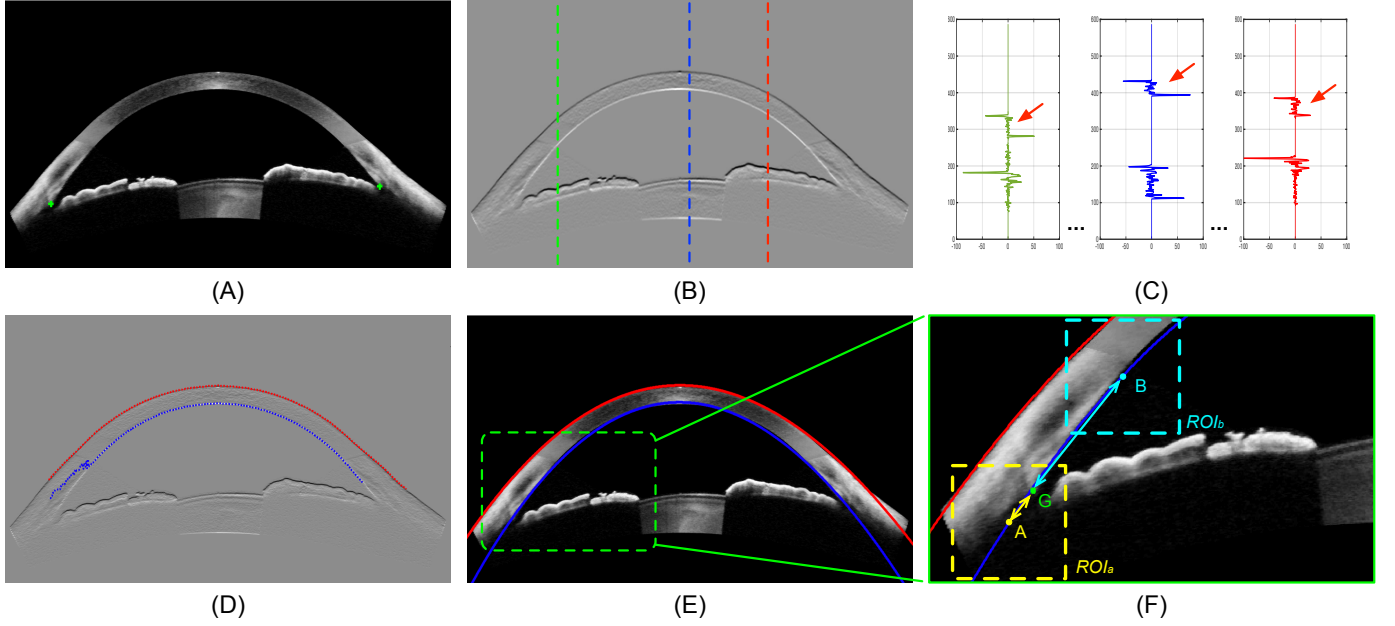
Fig. 4. (A) Original AS-OCT with the Scleral Spur points highlighted in green. (B) The vertical gradient map of an AS-OCT image. (C) The gradient value responses for the marked columns in the gradient map (B). The red arrow indicates the local impulse response corresponding to the corneal boundary. (D) The estimated initial points of the corneal boundary. (E) AS-OCT with the estimated corneal boundaries. The red and blue dotted curves denote the estimated corneal boundaries. (F) A zoom-in region of the AS-OCT image. Two sliding ROIs are marked by yellow and cyan rectangles. The distance between the ROI center (e.g., A, B) to the Scleral Spur point G is calculated and used for regression.

(SVR) [51]:

$$\min_{\mathbf{w}_r} \frac{1}{2}\mathbf{w}_r^T\mathbf{w}_r + C\sum_{i=1}^{l}(\max(0, |\mathbf{w}_r^T\mathbf{x}_i - d_i| - \epsilon)^2), \quad (2)$$

where $C > 0$ is a regularization parameter, $\mathbf{w}_r$ is the regression parameter vector, and l is the training number. The parameter ϵ is employed so that the loss is zero if $|\mathbf{w}_r^T\mathbf{x}_i - d_i| \leq \epsilon$. $(\mathbf{x}_i, d_i)$ are the ROI-label pairs with the feature vector $\mathbf{x}_i$ and corresponding distance d_i of ROI i. For ACA detection, we repeat the steps, and output the ROI with the smallest regression value as the ACA patch. In this paper, we implement the SVR model based on the liblinear toolbox [51], and the HOG [34] feature is utilized to represent the ROI.

B. Network architecture

Our proposed Multi-Level Deep Network (MLDN) consists of three parallel sub-networks to generate representations for multiple clinically-relevant regions in an AS-OCT image, as shown in Fig. 5, including the global anterior segment structure, local iris region, and ACA patch. Each sub-network contains a sequence of convolutional blocks composed of a convolutional layer, batch normalization [52], and ReLU activation [18]. In them, the convolutional layers learn and generate local feature maps based on local receptive fields over its input, while the batch normalization acts as a regularization that helps to avoid overfitting [52]. The ReLU is employed as the activation function following the convolutional layer.

The first sub-network of our MLDN works on the global AS-OCT image to represent the whole anterior segment structure. The anterior segment structure provides multiple clinical measurements [28], such as lens vault and anterior chamber structure. The second sub-network focuses on the local region containing the iris structure, as iris-related parameters (e.g., angle recess area, anterior chamber depth, iris shape and curvature) are widely used in clinical diagnosis of angle-closure [33]. This local region is obtained by cropping half of the AS-OCT image containing only one ACA region. The last sub-network processes a small patch centered on the ACA area. In clinical diagnosis, the ACA area is the most important because several major risk factors for angle-closure detection are calculated from it [6] (e.g., trabecular-iris angle and angle opening distance). For each sub-network, we resize the input images to 224×224 to implement other pre-trained deep models conveniently [18], [53].

The feature maps of the three sub-networks are combined and fed into one fully connected layer with softmax regression to predict the final detection result. As a binary classification task, the binary cross-entropy is utilized as the loss function for angle-closure detection:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i\log(p_i) + (1 - y_i)\log(1 - p_i)], \quad (3)$$

where p_i is estimated probability for the i-th image, $y_i \in \{0,1\}$ is the angle-closure label, and N denotes the number of training images.

C. Intensity-based data augmentation

The AS-OCT images are captured along the perpendicular direction of the eye, and anterior chamber structures generally appear at consistent positions in AS-OCT images. As a result,
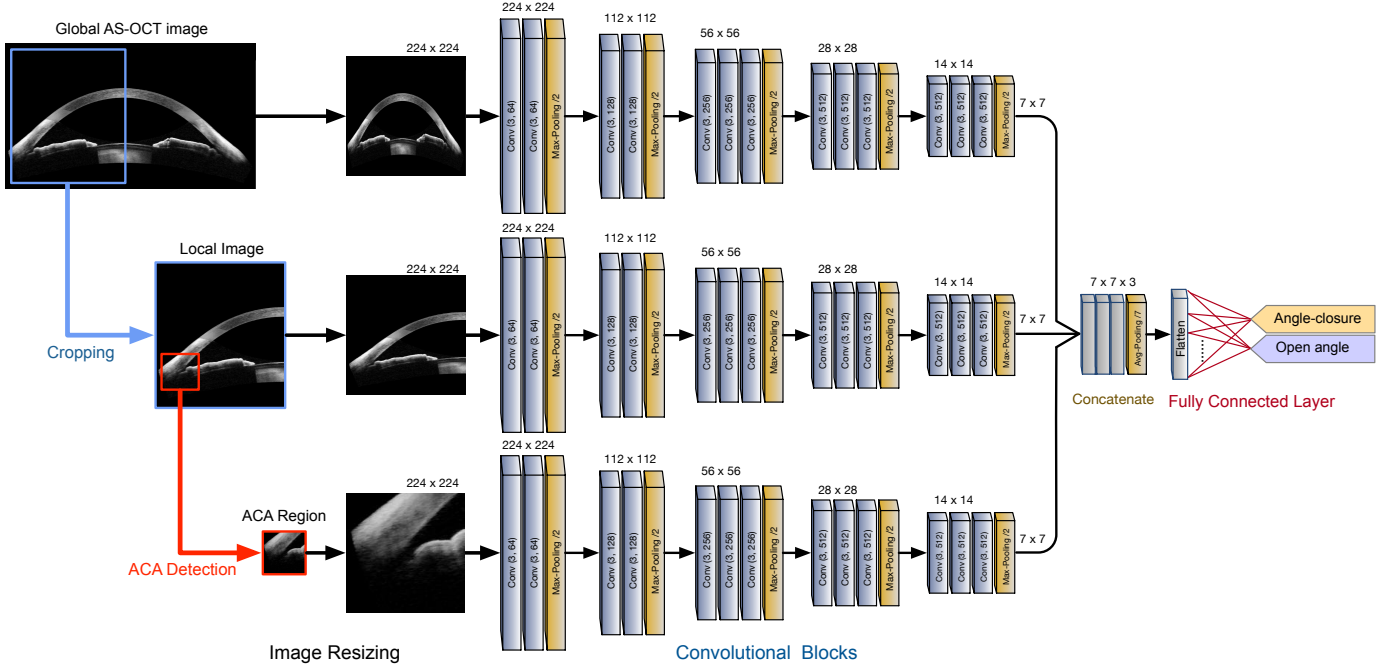
Fig. 5. Architecture of our multi-level deep network. Each convolutional block contains a convolutional layer, batch normalization, and ReLU activation. The merge layer (orange block) combines the feature maps of the three sub-networks, while one fully connected layer with softmax is used to predict the detection result. The parameters of the convolutional block are denoted as "Conv (channel number, kernel size)"

general image augmentation operations by rotation and scaling do not facilitate training for AS-OCT images. Moreover, the different AS-OCT modalities may lead to varied image intensities, which can degrade detection performance. Based on this, an intensity-based data augmentation is introduced on the training data, specifically by re-scaling image intensities by a factor of k_I ($k_I = [0.5, 1, 1.5]$ in this paper). We note that errors may exist in the localization of the ACA patch. To reduce the effect of such errors, we also augment the data by shifting the ACA center to generate different patches for input into the ACA sub-network.

IV. EXPERIMENTS

A. Clinical AS-OCT dataset

For experimentation, we collected two clinical AS-OCT datasets for different devices, namely a Carl Zeiss Visante AS-OCT and a Cirrus HD-OCT.

Visante AS-OCT dataset is captured using a Visante AS-OCT machine (Model 1000, Carl-Zeiss Meditec). The dataset is composed of 4135 AS-OCT images from 2113 subjects. As two ACA regions exist in one AS-OCT image, we split each AS-OCT into two ACA images (yielding a total of 8270 ACA images), and flip the right-side ACA images horizontally to align with the left-side images. The ground truth for each ACA image is labeled as open-angle or angle-closure. It is defined as the presence (angle-closure) or absence (open-angle) of iris-trabecular contact by visual inspection of the images by a clinician trained in examining the images, which are later adjudicated by a senior ophthalmologist trained and experienced in glaucoma diagnosis. Overall, the whole dataset contains 7375 open-angle and 895 angle-closure ACA images.
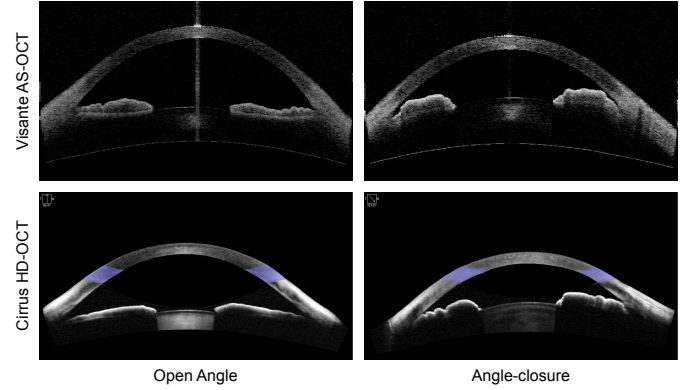

Fig. 6. Example images from the two AS-OCT datasets.

Cirrus HD-OCT dataset is captured using a Cirrus HD-OCT machine (Model HD 5000, Version 8.0, Carl-Zeiss Meditec, Germany). It contains 701 AS-OCT images in total from 202 subjects. Same as with Visante AS-OCT, two ACA images are extracted from each AS-OCT image, and finally we obtain 1102 open-angle and 300 angle-closure ACA images. Different from the Visante AS-OCT imaging, post-processing is done during Cirrus machine imaging, where the corneal boundary is made clearer, as shown in Fig. 6. However, the posterior border of the iris and the iris trabecular contact points are not obvious in all images.

In the experiment, each dataset is divided randomly into training and testing sets at a patient level, which ensures that the ACA images from one patient fall into the same training or test set.

B. Implementation

Training of the proposed MLDN is composed of two phases. First, the pre-trained model trained from ImageNet is loaded in each sub-network as the initialization, and then each sub-network is fine-tuned individually on the training dataset. Second, we combine the three trained sub-networks together into the MLDN and fine-tune the last fully connected layer on the training dataset. To optimize our deep models, stochastic gradient descent (SGD) is conducted on mini-batches with a learning rate of $1e-4$ and a momentum of 0.9. The whole framework is implemented by using Keras with Tensorflow. Fine-tuning (200 epochs) requires approximately five hours on one NVIDIA K40 GPU. In testing, the detection result is produced within 500 ms per AS-OCT image including ACA detection and classification.

C. Experimental criterion and baseline

For detection evaluation, Receiver Operating Characteristic (ROC) curves and area under ROC (AUC) curve by applying different thresholds are reported. Moreover, we also represent the highest point on the ROC curve as the diagnostic threshold to calculate four evaluation criteria, namely Sensitivity (Sen), Specificity (Spe), Balanced Accuracy ($BAcc$), and F-measure (F_m):

$$Sen = \frac{TP}{TP+FN}, \ Spe = \frac{TN}{TN+FP},$$
$$BAcc = \frac{1}{2}(Sen+Spe), \ F_m = \frac{2TP}{2TP+FP+FN},$$

where TP, TN, FP and FN denote the number of true positives, true negatives, false positives and false negatives, respectively.

We compare our algorithm with some angle-closure detection methods and deep learning networks:

- Clinical parameter based method [11], which segments the AS-OCT image by using the label transfer technique and calculates the pre-defined clinical parameters. The screening result is then obtained from the clinical parameters via a linear SVM.

- Visual feature based method [9], which detects the ACA region based on corneal structure and employs HOG features for classifying the angle-closure. The HOG feature is extracted from a 150×150 patch centered on the ACA region, and a linear SVM is used to predict the detection result.

- Single-network Deep Model. Three state-of-the-art deep learning architectures are used as baselines: VGG-16 [53], Inception-V3 [54], and ResNet50 [36]. VGG-16 has 13 convolutional layers and 3 fully connected layers [53]. Inception-V3 utilizes factorized smaller convolutions and aggressive spatial regularization to avoid representational bottlenecks [54]. ResNet50 employs a residual function to gain accuracy from considerably increased depth [36]. For each deep model, we report three results: one is based on the original global image (Img), the second is local iris region (Loc) as used in our second sub-network, and the last is based on the ACA patch (ACA) used in our third sub-network.

- Multi-Scale Deep Model: We also report results from two types of multi-scale deep learning models. The first is the multi-column deep model used in [44], [45] with different receptive field sizes for each sub-network, as shown in Fig. 2 (A). The second is the skip-layer deep model [46], [47], which combines different pooling layers to build the representation on a primary sub-network, as shown in Fig. 2 (B).

D. Detection evaluation

We test our MLDN on the Visante and Cirrus datasets. The performances are reported in Table I and Fig. 7. Most clinical measurements are defined according to anatomical structures, which are known to have specific physical significance in making a diagnosis. However, the clinical parameters do not represent information beyond what clinicians observed, and do account for detailed visual context of the AS-OCT. Thus the performance with clinical parameters does not perform best in automated screening based on AS-OCT. Visual representations could mine more image details, and is found to achieve better scores than the clinical parameter based method as expected. This was also observed in previous work [11].

The deep learning based methods can learn rich representations more powerful than clinical parameters and handcrafted visual features. In the Visante AS-OCT dataset, the deep models (e.g., VGG-16 and Inception-V3) perform better than visual features at all scales (Img, Loc, and ACA). But in the Cirrus HD-OCT dataset, only the performance of the VGG-16 model on Loc and ACA are higher than for visual features. A possible reason for this is that deeper architectures such as Inception-V3 and ResNet require larger-scale datasets for effective fine-tuning, while in our case only smaller-sized datasets (e.g., Cirrus HD-OCT dataset) are available. For different scales of inputs, the ACA patch focuses on the most highly relevant details, and in all the deep models, the results on the ACA patch are higher than for the global image or local iris region. For multi-scale deep models (Multi-Column and Skip-Layer models), the shallow layers can be easier to optimize, resulting in similar performance on both datasets. Our MLDN outperforms the other networks, showing that the proposed combination of multi-level clinically-relevant regions improves the detection performance, with AUC of 0.9619 (95% Confidence Interval (CI): 0.9499-0.9711) on the Visante AS-OCT dataset and 0.9524 (95% CI: 0.9208-0.9718) on the Cirrus HD-OCT dataset, respectively. Note that the AS-OCT images were captured with the standardized position and direction, which leads them to have similar structure locations. Alignment pre-processing is thus not needed. Moreover, the data augmentation of the deep learning method promotes robustness of our method against slight shifts and rotations.

E. Component evaluation

Our multi-level architecture employs three sub-networks to generate the screening result. To evaluate the effectiveness of each sub-network, we report the results on different combinations of both datasets, as shown in Table II and Fig. 8. From the results, we can see that the performance on global images is the worst. Single sub-networks on local regions and ACA patches obtain similar scores. Combinations of two sub-networks perform better than a single sub-network, and

TABLE I

PERFORMANCE OF DIFFERENT METHODS ON TWO DATASETS. THE BEST TWO RESULTS ARE SHOWN IN RED AND BLUE, RESPECTIVELY.

	Visante AS-OCT Dataset				Cirrus HD-OCT Dataset			
Methods	B-Accuracy	Sensitivity	Specificity	F-measure	B-Accuracy	Sensitivity	Specificity	F-measure
Quantitative Parameter	0.8186	0.7914	0.8459	0.5058	0.7824	0.8559	0.7090	0.5026
Visual Feature	0.8667	0.8776	0.8558	0.5609	0.8299	0.7748	0.8849	0.6491
VGG-16 Img	0.8823	0.8957	0.8689	0.5904	0.8465	0.7928	0.9002	0.6822
VGG-16 Loc	0.8895	0.8685	0.9104	0.6564	0.8357	0.8559	0.8156	0.6032
VGG-16 ACA	0.9003	0.8980	0.9026	0.6545	0.8792	0.8378	0.9205	0.7410
Inception-V3 Img	0.8771	0.8662	0.8880	0.6102	0.7698	0.7748	0.7648	0.5119
Inception-V3 Loc	0.8998	0.9070	0.8927	0.6390	0.7751	0.7838	0.7665	0.5179
Inception-V3 ACA	0.9008	0.8889	0.9128	0.6718	0.8690	0.8108	0.9272	0.7377
ResNet50 Img	0.8269	0.7664	0.8874	0.5587	0.7453	0.6937	0.7970	0.4968
ResNet50 Loc	0.8462	0.8209	0.8715	0.5591	0.7276	0.6937	0.7614	0.4681
ResNet50 ACA	0.8781	0.8481	0.9081	0.6410	0.7847	0.8198	0.7496	0.5200
Multi-Column Model	0.9056	0.9274	0.8838	0.6317	0.8595	0.7748	0.9442	0.7478
Skip-Layer Model	0.9039	0.8844	0.9235	0.6940	0.8602	0.8559	0.8646	0.6643
Our MLDN	0.9180	0.9297	0.9062	0.6777	0.9124	0.8739	0.9509	0.8186

TABLE II

PERFORMANCE OF DIFFERENT SUB-NETWORK COMBINATIONS ON TWO DATASETS.

	Visante AS-OCT Dataset				Cirrus HD-OCT Dataset			
Methods	B-Accuracy	Sensitivity	Specificity	F-measure	B-Accuracy	Sensitivity	Specificity	F-measure
MLDN (Img)	0.8969	0.8617	0.9321	0.7031	0.7786	0.7297	0.8274	0.5510
MLDN (Loc)	0.8958	0.8798	0.9117	0.6650	0.8292	0.8649	0.7936	0.5836
MLDN (ACA)	0.8932	0.8753	0.9112	0.6615	0.8476	0.7748	0.9205	0.7049
MLDN (Img + Loc)	0.9090	0.8934	0.9245	0.7011	0.8822	0.8829	0.8816	0.7025
MLDN (Img + ACA)	0.9090	0.8934	0.9245	0.7011	0.8732	0.8378	0.9086	0.7209
MLDN (Loc + ACA)	0.9028	0.9025	0.9031	0.6579	0.8941	0.8829	0.9052	0.7396
MLDN (All) w/o AG	0.9024	0.9116	0.8932	0.6422	0.8690	0.8919	0.8460	0.6578
MLDN (All)	0.9180	0.9297	0.9062	0.6777	0.9124	0.8739	0.9509	0.8186

our final architecture with all sub-networks obtains the best performance. Moreover, we report results obtained without the intensity-based augmentation (Our MLDN w/o AG). We can observe that the intensity-based augmentation yields clear improvements, with the AUC score increasing from 0.950 to 0.962 on the Visante AS-OCT dataset and from 0.937 to 0.952 on the Cirrus HD-OCT dataset.

V. CONCLUSION

In this paper, we presented an automated angle-closure detection system for AS-OCT images. A multi-level deep network was designed to obtain multi-level AS-OCT representations on clinically meaningful regions. In addition, an intensity-based augmentation for AS-OCT was introduced in the training phase. Experiments on two clinical AS-OCT datasets demonstrate that the proposed MLDN method outperforms previous detection methods and other deep learning networks. The design of MLDN was guided by clinical prior knowledge, and the intensity-based augmentation could also be utilized in other OCT-related tasks. Moreover, our MLDN architecture is a general framework not limited to AS-OCT images. Actually, it can be used in any real-life application that would benefit from extraction of a meaningful foreground region and integrating it with global context. In the future, we will examine further applications of MLDN.

REFERENCES

[1] H. Quigley and A. Broman, "The number of people with glaucoma worldwide in 2010 and 2020," *Br. J. Ophthalmol.*, vol. 90, no. 3, pp. 262–267, 2006.

[2] Y.-C. Tham, X. Li, T. Y. Wong, and et al, "Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis," *Ophthalmology*, vol. 121, no. 11, pp. 2081–2090, 2014.

[3] R. Sharma, A. Sharma, T. Arora, and et al, "Application of anterior segment optical coherence tomography in glaucoma," *Surv. Ophthalmol.*, vol. 59, no. 3, pp. 311–327, 2014.

[4] M. Ang, M. Baskaran, R. M. Werkmeister, and et al, "Anterior segment optical coherence tomography," *Progress in Retinal and Eye Research*, 2018.

[5] H. T. Wong, M. C. Lim, L. M. Sakata, and et al, "High-definition optical coherence tomography imaging of the iridocorneal angle of the eye," *Archives of Ophthalmology*, vol. 127, no. 3, pp. 256–260, 2009.

[6] M. Nongpiur, B. Haaland, D. Friedman, and et al, "Classification algorithms based on anterior segment optical coherence tomography measurements for detection of angle closure," *Ophthalmology*, vol. 120, no. 1, pp. 48–54, 2013.

[7] M. E. Nongpiur, I. F. Aboobakar, M. Baskaran, and et al, "Association of Baseline Anterior Segment Parameters With the Development of Incident Gonioscopic Angle Closure," *JAMA Ophthalmology*, vol. 135, no. 3, pp. 252–258, 2017.

[8] J. Tian, P. Marziliano, M. Baskaran, H. Wong, and T. Aung, "Automatic anterior chamber angle assessment for HD-OCT images," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 11, pp. 3242–3249, 2011.

[9] Y. Xu, J. Liu, N. Tan, and et al, "Anterior chamber angle classification using multiscale histograms of oriented gradients for glaucoma subtype identification," in *EMBC*, 2012, pp. 3167–3170.

[10] Y. Xu, J. Liu, J. Cheng, and et al, "Automated anterior chamber angle localization and glaucoma type classification in OCT images," in *EMBC*, 2013, pp. 7380–7383.
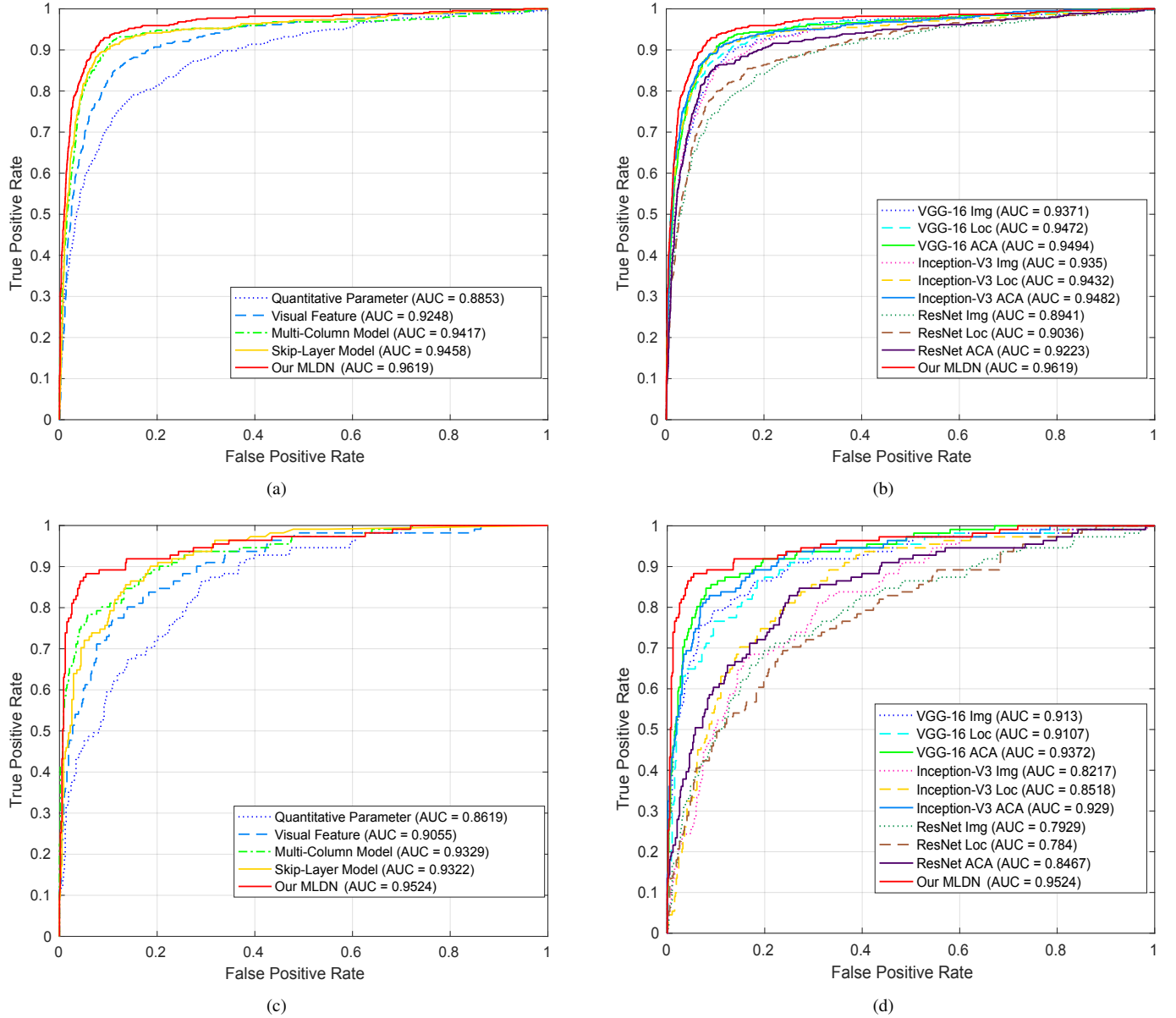
Fig. 7. Performance of different methods on the (A)-(B) Visante AS-OCT dataset and (C)-(D) Cirrus HD-OCT dataset.

[11] H. Fu, Y. Xu, S. Lin, and et al, "Segmentation and Quantification for Angle-Closure Glaucoma Assessment in Anterior Segment OCT," *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1930–1938, 2017.

[12] ——, "Multi-Context Deep Network for Angle-Closure Glaucoma Screening in Anterior Segment OCT," in *International Conference On Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2018, pp. 356–363.

[13] J. Console, L. Sakata, T. Aung, D. Friedman, and M. He, "Quantitative analysis of anterior segment optical coherence tomography images: the zhongshan angle assessment program," *Br. J. Ophthalmol.*, vol. 92, no. 12, pp. 1612–1616, 2008.

[14] H. Fu, Y. Xu, D. Wong, and et al, "Automatic Anterior Chamber Angle Structure Segmentation in AS-OCT Image based on Label Transfer," in *EMBC*, 2016, pp. 1288–1291.

[15] L. Deng and D. Yu, "Deep Learning: Methods and Applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 34, pp. 197–387, 2014.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[17] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[18] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *NIPS*, 2012, pp. 1097–1105.

[19] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep Learning for Generic Object Detection: A Survey," *arXiv*, 2018.

[20] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2017.

[21] V. Gulshan, L. Peng, M. Coram, and et al, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *Journal of the American Medical Association*, vol. 304, no. 6, pp. 649–656, 2016.

[22] D. Ting, C. Cheung, G. Lim, and et al, "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes," *JAMA*, vol. 318, no. 22, pp. 2211–2223, 2017.

[23] Y. Zhao, Y. Zheng, Y. Liu, and et al, "Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 51–63, 2017.

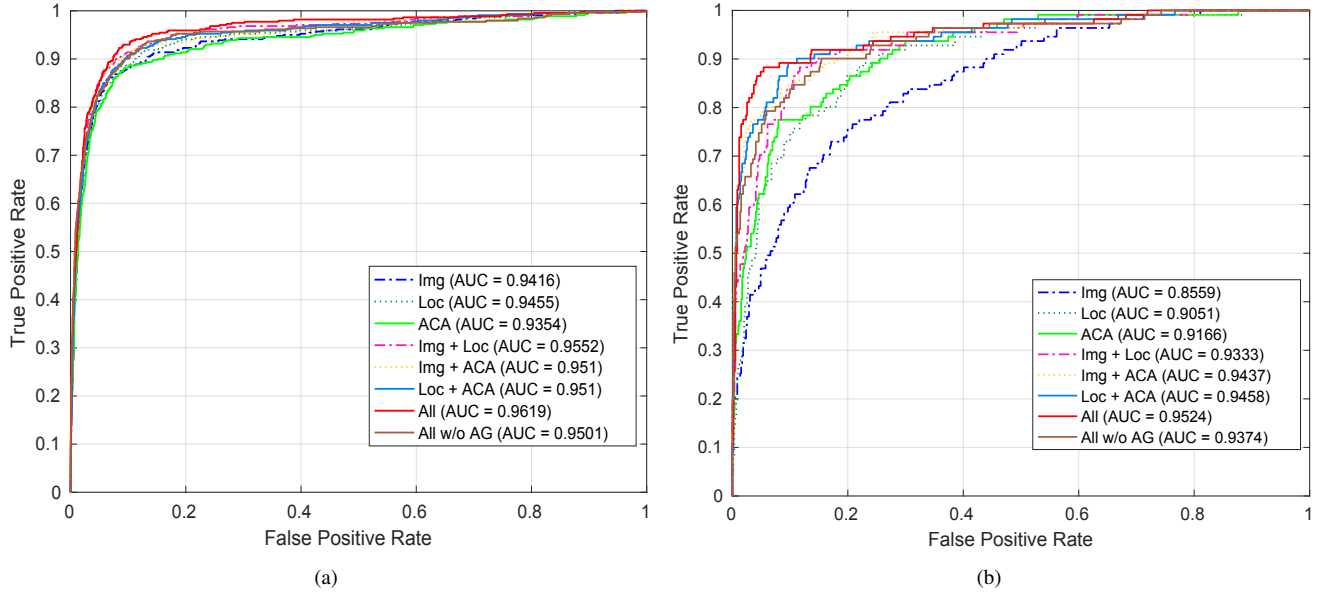[24] F. Grassmann, J. Mengelkamp, C. Brandl, and et al, "A Deep Learning

Fig. 8. Performance of different sub-network combinations in our MLDN on the (A) Visante AS-OCT and (B) Cirrus HD-OCT datasets.

Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography," *Ophthalmology*, pp. 1–11, 2018.

[25] H. Liu, D. Wong, H. Fu, Y. Xu, and J. Liu, "DeepAMD: Detect Early Age-Related Macular Degeneration by Applying Deep Learning in a Multiple Instance Learning Framework," in *ACCV*, 2018.

[26] H. Fu, J. Cheng, Y. Xu, and et al, "Joint Optic Disc and Cup Segmentation Based on Multi-label Deep Network and Polar Transformation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.

[27] ——, "Disc-aware Ensemble Network for Glaucoma Screening from Fundus Image," *IEEE Transactions on Medical Imaging*, 2018.

[28] R. Wu, M. Nongpiur, M. He, and et al, "Association of narrow angles with anterior chamber area and volume measured with anterior-segment optical coherence tomography," *Arch. Ophthalmol.*, vol. 129, no. 5, pp. 569–574, 2011.

[29] D. Williams, Y. Zheng, F. Bao, and A. Elsheikh, "Automatic segmentation of anterior segment optical coherence tomography images," *J. Biomed. Opt.*, vol. 18, no. 5, pp. 1–7, 2013.

[30] S. Niwas, W. Lin, C. Kwoh, and et al, "Cross-examination for angle-closure glaucoma feature detection," *IEEE J. Biomed. Health. Inf.*, vol. 20, no. 1, pp. 343–354, 2016.

[31] C. Leung, H. Li, R. Weinreb, and et al, "Anterior Chamber Angle Measurement with Anterior Segment Optical Coherence Tomography: A Comparison between Slit Lamp OCT and Visante OCT," *Invest. Ophthalmol. Vis. Sci.*, vol. 49, no. 8, pp. 3469–3474, 2008.

[32] G. Tan, M. He, W. Zhao, and et al, "Determinants of Lens Vault and Association With Narrow Angles in Patients From Singapore," *Am. J. Ophthalmol.*, vol. 154, no. 1, pp. 39–46, 2012.

[33] S. Ni Ni, J. Tian, P. Marziliano, and H. Wong, "Anterior Chamber Angle Shape Analysis and Classification of Glaucoma in SS-OCT Images," *Journal of Ophthalmology*, pp. 1–12, 2014.

[34] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *CVPR*, vol. 1, 2005, pp. 886–893.

[35] M. J. Gangeh, L. Sørensen, S. B. Shaker, and et al, "A Texton-Based Approach for the Classification of Lung Parenchyma in CT Images," in *MICCAI*, 2010, pp. 595–602.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *CVPR*, 2016, pp. 770–778.

[37] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in *CVPR*, 2014.

[38] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[39] H. Zhu, R. Vial, S. Lu, and et al, "YouTube: Searching Action Proposal via Recurrent and Static Regression Networks," *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 2609–2622, 2018.

[40] S. Zheng, S. Jayasumana, B. Romera-Paredes, and et al, "Conditional Random Fields as Recurrent Neural Networks," in *ICCV*, 2015.

[41] H. Fu, Y. Xu, S. Lin, and et al, "DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional Random Field," in *MICCAI*, 2016, pp. 132–139.

[42] L. Zheng, Y. Yang, and Q. Tian, "SIFT Meets CNN: A Decade Survey of Instance Retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 5, pp. 1224–1244, 2018.

[43] J. Lei, L. Niu, H. Fu, and et al, "Person Re-Identification by Semantic Region Representation and Topology Constraint," *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, In press.

[44] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *CVPR*, 2012, pp. 3642–3649.

[45] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-Image Crowd Counting via Multi-Column Convolutional Neural Network," in *CVPR*, 2016, pp. 589–597.

[46] D. Eigen, C. Puhrsch, and R. Fergus, "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network," in *NIPS*, 2014, pp. 2366–2374.

[47] S. Yang and D. Ramanan, "Multi-scale recognition with DAG-CNNs," in *CVPR*, 2016, pp. 1215–1223.

[48] G. Li and Y. Yu, "Visual Saliency Based on Multiscale Deep Features," in *CVPR*, 2015, pp. 5455–5463.

[49] J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders, "Selective Search for Object Recognition," *Int. J. Comput. Vision*, vol. 104, no. 2, pp. 154–171, 2013.

[50] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, 2011.

[51] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A Library for Large Linear Classification," *Journal of Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.

[52] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *ICML*, 2015, pp. 448–456.

[53] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*, 2015.

[54] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *CVPR*, 2016, pp. 2818–2826.